Information Fusion 49 (2019) 89–99

Contents lists available at ScienceDirect

Information Fusion

journal homepage: www.elsevier.com/locate/inffus

Deep learning with multi-scale feature fusion in remote sensing for automatic oceanic eddy detection

Yanling Du[a,b], Wei Song[a,*], Qi He[a], Dongmei Huang[a], Antonio Liotta[c], Chen Su[b]

[a] College of Information and Technology, Shanghai Ocean University, Shanghai 201306, China
[b] East China Sea Forecast Center of State Oceanic Administration, Shanghai 200136, China
[c] University of Derby, Derby DE1 3HD, UK

ARTICLE INFO

Keywords:
Remote sensing
Feature fusion
SAR images
Eddy detection
Deep learning

ABSTRACT

Oceanic eddies are ubiquitous in global oceans and play a major role in ocean energy transfer and nutrients distribution, thus being significant for understanding ocean current circulation and marine climate change. They are characterized by a combination of high-speed vertical rotations and horizontal movements, leading to irregular three-dimensional spiral structures. While the ability to detect eddies automatically and remotely is crucial to monitoring important spatial–temporal dynamics, existing methods are inaccurate because eddies are highly dynamic and the underlying physical processes are not well understood. Typically, remote sensing is used to detect eddies based on physical parameters, geometrics or other handcrafted features. In this paper, we show how Deep Learning may be used to reliably extract higher-level features and then fuse multi-scale features to identify eddies, regardless of their structures and scales. We learn eddy features using two principal component analysis convolutional layers, then perform a non-linear transformation of the features through a binary hashing layer and block-wise histograms. To handle the difficult problem of spatial variability across synthetic aperture radar (SAR) images, we introduce a spatial pyramid model to allow multi-scale features fusion. Finally, a linear support vector machine classifier recognizes the eddies. Our method, dubbed DeepEddy, is benchmarked against a dataset of 20,000 SAR image samples, achieving a 97.8 ± 1% accuracy of detection.

1. Introduction

Object detection is one of the most active research topics in the field of computer vision, because it is the basis for target identification, monitoring and tracking [1–3]. Automatic remote sensing image detection is crucial for monitoring the dynamics of oceanic eddies, which are ubiquitous in global oceans and dominate more than half of the ocean's kinetic energy [4,5]. The dynamic nature of oceanic eddies is caused by the instability of ocean currents, the stress of sea surface wind, and the variety of seabed topography [6]. As a result, oceanic eddies have rotating coherent structures in diverse patterns, with diameters varying between several hundred meters to several hundred kilometers. Conventionally, eddies having a radius of 10–100 km are referred to as 'mesoscale eddies', while those with a radius shorter than 10 km are called 'sub-mesoscale eddies'. What makes eddies particularly difficult to detect is that they move rapidly and dynamically change their structures, velocity and vorticity, during a lifetime that can span from days to hundreds of days [7].

Oceanic eddy detection based on the variety of remote sensing data has been widely studied [7–32]. However, great differences between the current oceanic eddy detection methods are shown, mainly because the information features of oceanic eddies acquired by different remote sensing satellites vary greatly. For instance, the changes in sea surface temperature (SST), chlorophyll concentration (CHL), sea surface height (SSH) and roughness caused by oceanic eddies can be recorded by infrared, ocean color, altimeter and synthetic aperture radar remote sensing, respectively. Eddy detection using SST and SSH is prone to false positives because many other ocean phenomena also impact the sea surface temperature and surface ocean color (chlorophyll concentration) [8], and they cannot detect the oceanic eddies with a diameter smaller than 50 km, because of the low spatial resolution of SSH data [7].

Synthetic aperture radar (SAR) images acquired oversea areas (Fig. 1), contain a wide range of small-scale and mesoscale phenomena of the ocean and are less susceptible to cloud and sunlight conditions [9]. Therefore, SAR images become an ideal data source to detect the oceanic eddies [10]. However, the oceanic eddies in SAR images manifest not only a variety of spatial scales and spiral structures (intensity, diameter, center) but also geometric deformation caused by the

* Corresponding author.
E-mail addresses: wsong@shou.edu.cn (W. Song), qihe@shou.edu.cn (Q. He), dmhuang@shou.edu.cn (D. Huang), a.liotta@derby.ac.uk (A. Liotta), sucheng@eastsea.gov.cn (C. Su).

https://doi.org/10.1016/j.inffus.2018.09.006
Received 28 June 2018; Accepted 9 September 2018
Available online 13 September 2018
1566-2535/© 2018 Elsevier B.V. All rights reserved.

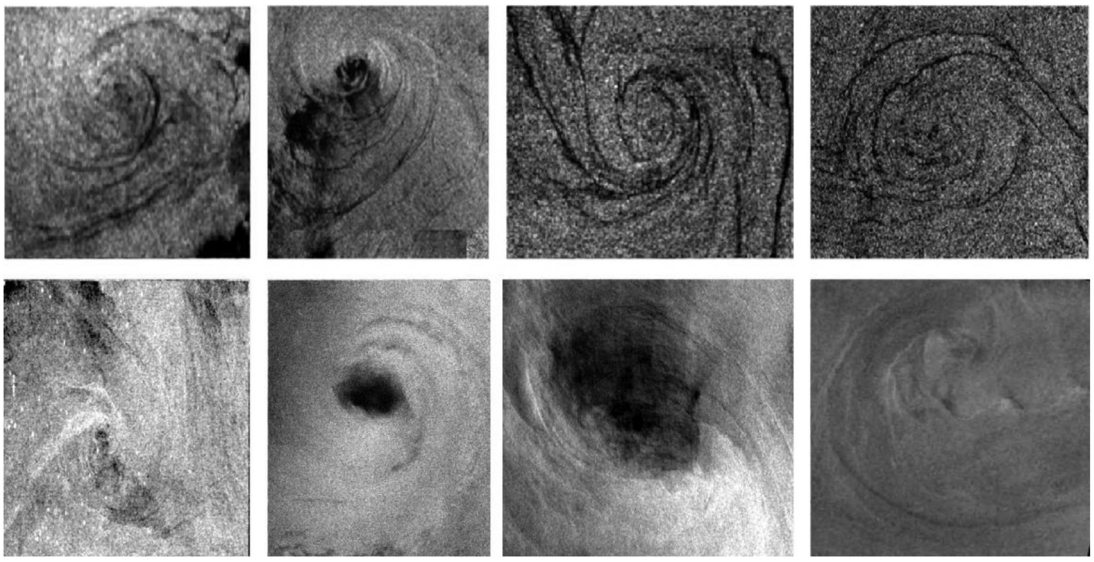

Fig. 1. Oceanic eddies in SAR images from multiple satellite sensors.

imaging mode of remote sensing satellites. In addition, some eddies in SAR imagery are dark (i.e., low backscatter), narrow, curvilinear, concentric bands that appeared to spiral inward. Whereas the other eddies (e.g., many were found in the Gulf Stream and East Australia) are a narrow band of enhanced brightness [32]. All of these features make oceanic eddy detection a challenging task.

The conventional methods of eddy detection are mainly divided into the following categories: physical parameter-based [7,10,14–27], geometric-based [11–13] or handcrafted feature-based [29–31]. Physical parameter-based methods need to set a suitable threshold for a specific region, and are therefore task-specific and limited in generalization capability. Geometric-based methods do not easily detect eddy details from the data without clear geometrical features. Handcrafted feature-based methods usually design low-level features by experts, which are not able to capture the full dynamics of oceanic eddies. All current methods are highly depended on much expert knowledge and are very sensitive to data noises. What is really missing is a generalized method to automatically detect oceanic eddies from multisource remotely sensed data, which is the subject of our paper.

Oceanic eddies are mostly active in the areas where ocean currents exist. For example, in the Kuroshio extension regions (in English "Japanese Current", i.e. a north-flowing ocean current on the west side of the North Pacific Ocean), cyclonic/anticyclonic eddies may cause local upwelling/downwelling, leaving footprints in the SST field in the form of local extremes. Therefore, geometric-based methods, such as geometric features of the velocity field [12] and quasi-contour tracing and clustering [13], have been widely used to identify the eddy centers, size, intensity, tracks and lifetime based on the SST and ocean color remotely sensed data. However, the significant differences in structures, scales, centers and other attributes between different ocean eddies lead to low detection accuracy of the geometric-based methods.

Observations from satellite radar altimeters provide sea surface height (SSH) maps, showing closed-contour sea level anomaly (SLA) of the oceanic eddies. Specifically, the anticyclonic eddies show positive SLAs, while the cyclonic eddies show negative SLAs. However, oceanic eddy detection methods on SSH maps have mainly been based on physical criteria, or the Okubo–Weiss (W) parameter method [7,14–18]. This is commonly used for global ocean eddies detection and monitoring, but has significant limitations due to its dependency on region-specific thresholds and sensitivity to noise in the SSH maps [19–21]. Alternative physical parameter-based oceanic eddies detection methods, such as winding angle [22,23] and wavelet analysis [24,25], also suffer from the dependency on velocity fields and are sensitive to the amplified noise in the derivatives of SSH data. Moreover, the pixel-based methods have poor computational performance when concerning the high-resolution remotely sensed data [19,20].

To resolve the weakness of the physical parameter-based methods, contour-based methods identify eddies globally as closed contoured smoothed SSH anomalies using a nearest neighbor search [19]. An improved algorithm adapted from PDELTA [26] could identify multiple regions gradually within a time-series, by adding a spatial analysis feature, and greatly raised the computational performance by visiting each pixel at most once [27]. Unfortunately, the SSH data have a spatial resolution of tens of kilometers and, thus, are not adequate to perform detailed analysis of oceanic eddies with a smaller diameter, such as the sub-mesoscale oceanic eddies [10].

SAR images obtained from radar sensors have a high spatial resolution and a wide swath of observation, which makes them effective sources for gaining more comprehensive and detailed information on mesoscale and sub-mesoscale phenomena in the oceans [9,28,29]. Currently, most studies of oceanic eddies detection in SAR images are based on handcrafted features [28–31]. The feature design requires expert knowledge and is not efficient with respect to time and labor, especially considering the massive and continuous producing of SAR images by the global earth observation system. In a recent research, a preliminary study on ocean eddy identification and tracking based on SLA data was proposed by using neural networks. However, the network training data was based on the results of Okubo–Weiss methods and the identification accuracy was unknown [33]. A hierarchical network with fuzzy logic rule-based classification was proposed to detect the ocean mesoscale atmospheric phenomena automatically, and it reached an accuracy of 80% [9]. These works indicate the great potential of hierarchical networks in ocean eddies detection.

In contrast to designing features based on expert knowledge, deep learning models can learn high-level features automatically from a large database. In the field of general object detection and target recognition from images, such as face recognition, facial expression recognition, and handwriting recognition, deep learning models have achieved great success [1,3,34–37]. Especially for some detection tasks with a complex background or changing environment, the feature fusion method based on deep learning is more effective [3]. However, the deep learning models need to train a huge number of parameters with a rich training dataset. A simple deep learning baseline network named PCANet [35] was proposed to reduce the computation expenses. The PCANet

network was designed to be easy to train and be adaptable to different tasks. Its performance was on par with the state-of-the-art deep learning models for different tasks such as image classification, object detection and pattern recognition [35]. Moreover, the principal component analysis (PCA) method has been a popular choice for remotely sensed image fusion [38].

Inspired by the general deep learning models and PCANet for object detection and image classification, we aimed to develop an effective and generalized method to solve the problem of oceanic eddy detection using SAR image data. The proposed automatic method, dubbed DeepEddy is based on PCA filters convolution to form a simple deep learning network, which can obtain invariant and high-level oceanic eddies features by layer-wise learning from the training dataset. As pyramid representation for image fusion can improve classification accuracy [39], DeepEddy adopts spatial pyramid pooling (SPP) in feature-level fusion. The SPP can handle the complicated structures and poses of oceanic eddies via fusing the multi-scale features, so that the trained DeepEddy method can be applied to detect diverse geometrical shapes of oceanic eddies in different scales and for different SAR image resolutions. DeepEddy, is benchmarked against a dataset of 20,000 SAR image samples, achieves a $97.8 \pm 1\%$ accuracy, which is significantly higher than the 80% reported in the automatic classification method for mesoscale atmospheric phenomena above sea [9].

The main contributions of our work are as follows: (1) A new method for oceanic eddy detection based on deep learning is proposed, which can hierarchically learn high level and invariant features of oceanic eddies. (2) An eddy dataset based on SAR images from multiple satellite sensors is first established for eddy detection, which provides a data benchmark for future research on eddy detection. Our dataset is publically available at https://github.com/YanlingDu-SHOU/Mesoscale-Oceanic-Eddy-Dataset. (3) The method achieves high accuracy for eddy detection without any manual intervention and expert knowledge, which puts forward a new idea for the accurate identification of high dynamic ocean phenomena. Additionally, it is also possible to detect small-scale oceanic eddies with a diameter of less than 50 km, which would not be possible with SSH data.

This paper is a substantial extension from our initial findings published in the ICNSC 2017 conference [37]. We give here all the details of DeepEddy, providing a comprehensive evaluation under a broader range of conditions and a comparative evaluation between DeepEddy and PCANet. Moreover, we benchmark our results against the visual interpretation method in SAR images presented in [32].

The rest of the paper is organized as follows. In the next section, we describe the method of oceanic eddy detection. In particular, we give more details about the spatial pyramid pooling (SPP) layer, which plays a key role in enhancing the DeepEddy robustness on dynamic oceanic eddy detection. Section 3 contains experiments and result analysis. We describe the established process of the oceanic eddy dataset and add the method of data augmentation. Then, the experiments under different parameters are performed and comprehensive results analysis is presented. In Section 4, we carry out comparisons between DeepEddy and PCANet in terms of the detection accuracy, and between DeepEddy and the visual interpretation method in SAR images. Finally, we conclude by discussing the implications of the presented study and future work in Section 5.

2. DeepEddy method

In order to detect oceanic eddies in SAR images automatically, we propose a method based on deep learning that can detect oceanic eddies within a range of spatiotemporal situations, without requiring expert knowledge. The details are put forward in the following sections.

2.1. A deep learning architecture for oceanic eddy detection

The software architecture of our proposed method, consisting of hierarchical feature learning layers and a linear support vector machine (SVM) classifier, is illustrated in Fig. 2. Firstly, the convolution filters used in DeepEddy are learned through principal component analysis (PCA) based on the vectorized representation of input SAR images. Local patterns of the eddy are then extracted in the PCA convolutional layer. The convolutional filter is an essential parameter in convolutional neural networks (CNNs) that directly determines the features of learning. Unlike the traditional CNNs, which initialize the convolution filters in a random way, DeepEddy uses the primary convolution filters that are generated from the principal component analysis to learn much more distinguishing features with a simple architecture. Although multiple convolution layers with PCA filters can be applied to learn high-level features of the eddy, our method only uses two PCA convolutional layers because experiments showed that two layers were sufficient to achieve high detection accuracy.

Next, a nonlinear processing layer that contains binarization and hashing mapping enhances the separation of the features. Then, the feature pooling layer with block-wise histograms reduces the dimensions of the features. Meanwhile, considering the dynamics of eddy shapes and poses caused by the phases and incident angles of the remote satellite sensors, spatial pyramid pooling (SPP) is used to extract multi-scale spatial features of oceanic eddies. Finally, based on the learned high-level and invariant features, the SVM classifier outputs the detection results.

Suppose that the training SAR images set $\{I_i\}_{i=1}^{N}$ contains N input images. Firstly, we resize each image to the same size $m \times n$. The patch size is $k_1 \times k_2$ at all convolutional layers. The parameters need to learn from the training SAR images. $\{I_i\}_{i=1}^{N}$ are only the PCA filters. The components of the proposed architecture are described in detail next.

2.2. The first convolution layer

In order to obtain the PCA filters of the convolution layer, we first need to vectorize the input SAR images. For each SAR image I_i in the training image set, a $k_1 \times k_2$ patch slides on the images with a certain stride. Then, we subtract the patch mean from each patch and get $\overline{X}_l = [\bar{x}_{i,1}, \bar{x}_{i,2}, \ldots, \bar{x}_{\tilde{m},\tilde{n}}]$, where $\tilde{m} = m - \lceil k_1/2 \rceil$, $\tilde{n} = n - \lceil k_2/2 \rceil$, $\bar{x}_{i,j} = x_{i,j} - \frac{1^T x_{i,j}}{k_1 k_2} 1$ is a mean-removed patch, and 1 is an all-one vector with proper dimensions. For all input images, we combine the vectors as (1).

$$X = \left[\bar{X}_1, \bar{X}_2, \ldots, \bar{X}_N\right] \in R^{k_1 k_2 \times N \tilde{m} \tilde{n}} \tag{1}$$

The filter number in the ith convolutional layer is denoted as L_i, and we use PCA to minimize the reconstruction error in a set of orthonormal filters, shown as (2).

$$\min_{V \in R^{k_1 k_2 \times L_1}} \left\| X - VV^{\mathrm{T}} X \right\|_F^2, \quad \text{s.t.} \quad VV^T = I_{L_1} \tag{2}$$

where I_{L_1} is a $L_1 \times L_1$ identify matrix and L_1 is the principle eigenvectors of XX^{T}.

The PCA filters can be expressed as (3).

$$W_l^1 = \mathrm{mat}_{k_1,k_2}\left(q_l\left(XX^{\mathrm{T}}\right)\right) \in R^{k_1 k_2}, \quad l = 1, 2, \ldots, L_1 \tag{3}$$

where the function $\mathrm{mat}_{k_1,k_2}(\mathrm{vector})$ is to convert the $\mathrm{vector} \in R^{k_1 \times k_2}$ to a matrix $W \in R^{k_1 \times k_2}$ and $q_l(XX^{\mathrm{T}})$ represents the lth principal eigenvector of XX^{T}. The dominating principal eigenvectors capture the main variation of all of the mean-removed training patches. The output of the first convolutional layer is defined as

$$I_i^l = I_i * W_l^1, \quad i = 1, 2, \ldots, N \tag{4}$$

where $*$ denotes the convolution of two dimensions. I_i^l is the ith input image, and W_l^1 is the lth filter of the PCA in the first convolutional layer.

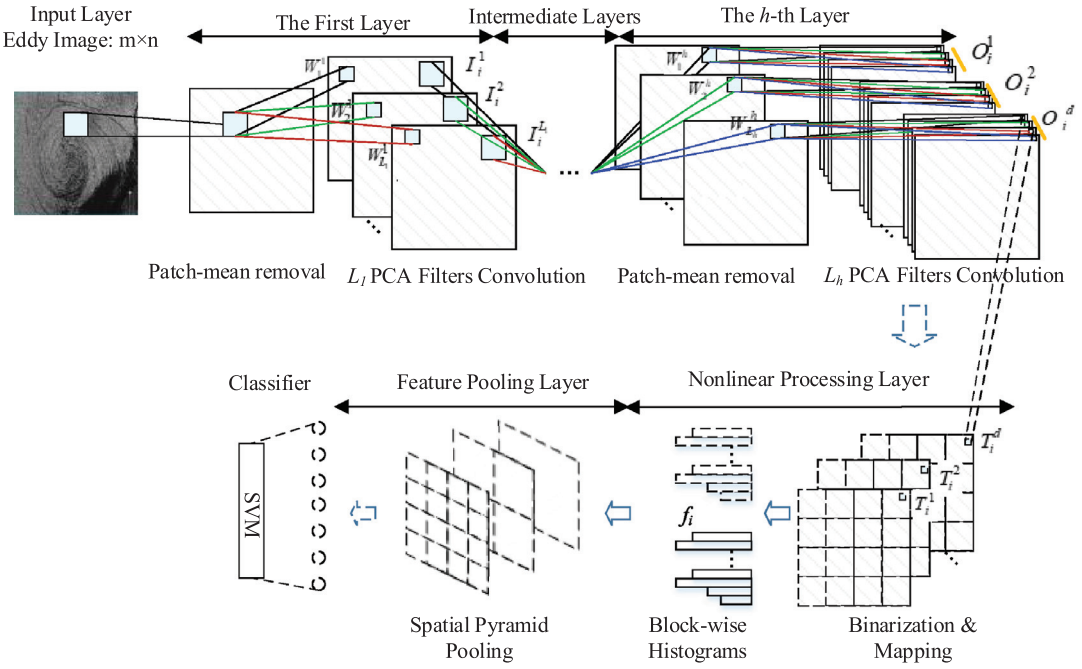

Fig. 2. DeepEddy, the method based on a simple deep architecture for oceanic eddy detection. The method mainly consists of PCA-based convolution layers, a nonlinear processing layer and a feature pooling layer to learn the high level and invariant features of oceanic eddies.

2.3. The second convolution layer

A majority of operations in the second and above layers are similar to the first convolutional layer. We subtract the patch mean from each patch and collect the mean-removed patches to concatenate them into:

$$Y^l = \left[\bar{Y}_1^l, \bar{Y}_2^l, \ldots, \bar{Y}_N^l\right] \in R^{k_1 k_2 \times N\tilde{m}\tilde{n}} \tag{5}$$

where Y^l represents the outputs of all the images after convolving with W_l^1.We can obtain the PCA filters of the higher layers, which are denoted as

$$W_l^h = \mathrm{mat}_{k_1,k_2}\left(q_l\left(YY^T\right)\right) \in R^{k_1 \times k_2}, \quad l = 1, 2, \ldots, L_h, h \geq 2 \tag{6}$$

For each input I_i^l of the layer h convolving with W_l^h for $l = 1, 2, \ldots, L_h$, the output O will have L_h images of the size $m \times n$.

$$O_i^l = \left\{I_i^l * W_l^h\right\}_{l=1}^{L_h} \tag{7}$$

The architecture can be easily built with more layers if a deeper architecture is found to be beneficial. At the hth layer, D ($\boldsymbol{D} = \Pi_{(i=1)}^{(h-1)} L_i$) inputs produces D sets of outputs $O_i^d (d = 1, 2, \ldots, D)$ and each set has L_h images. Thus, the total number of output images will be $\Pi_{i=1}^{h} L_i$.

2.4. Nonlinear processing layer

After the convolution layers, the feature maps of the input SAR images are obtained. We use hashing to operate nonlinear processing on the convolutional outputs. The hth layer outputs have D ($D = \prod_{i=1}^{h-1} L_i$) sets of real values and each set has a total of L_h outputs $\{I_i^l * W_l^h\}_{l=1}^{L_h}$. The outputs are binarized to $\{H(I_i^l * W_l^h)\}_{l=1}^{L_h}$ using a Heaviside step function $H(\cdot)$, whose value is one for positive entries and zero otherwise. Following binarization, each binary string is encoded into decimal numbers, and the map function is as follows:

$$T_i^d = \sum_{l=1}^{L_h} 2^{l-1} \mathrm{H}\left(O_i^d\right), \quad d = 1, 2, \ldots, D \tag{8}$$

This forms an "image" for each set of L_h output, where each pixel value is an integer and ranges from 0 to $2^{L_h} - 1$.

2.5. Feature pooling layer

The feature pooling is performed in two steps: the block-based histogram is used to reduce feature dimensions. Then spatial pyramid pooling is used to extract the multi-scale features of oceanic eddies to improve the robustness of DeepEddy.

2.5.1. Block-based histogram

After the nonlinear processing, each of these "image" T_i^l is partitioned into B blocks. We compute the histogram of the decimal values in each block and concatenate all B histograms into one vector and denote this vector as $\mathrm{Bhist}(T_i^l)$. Then, the "features" of the input SAR image I_i is defined as the set of block-wise histograms, i.e.,

$$\boldsymbol{f}_i \doteq \left[\mathbf{Bhist}\left(\boldsymbol{T}_i^1\right), \ldots, \mathbf{Bhist}\left(\boldsymbol{T}_i^D\right)\right]^T \in \boldsymbol{R}^{\left(2^{L_h}\right)\cdot \boldsymbol{D}\cdot \boldsymbol{B}} \tag{9}$$

In this process, the local blocks are overlapped, and the overlapping ratio is another important parameter of DeepEddy, as will be further discussed in Section 3. The block-wise histogram encodes special information and offers some degree of translation invariance in the learned features within each blocks.

2.5.2. Spatial pyramid pooling

Oceanic eddies captured in SAR images have significant geometric distortion and spatial scale differences. This is caused by the spatial and temporal dynamic nature of oceanic eddies as well as the multi-temporal, multi-angle, multi-polarization imaging mode of SAR satellites. To deal with the complexity of oceanic eddies in structures and scales, the spatial pyramid pooling (SPP) method is introduced into the DeepEddy architecture, as the flowchart shows in Fig. 3.

The features can map to a fixed length feature vector in the feature pooling layer based on the multiscale partition. Specifically, the length of the feature vector is 21 (16 + 4 + 1) in our model. Assuming that the sizes of the input feature maps are (w, h), the steps of SPP are as follows.

(1) In the partition based on the first scale, one input feature map is divided into 16 sub-image blocks with a size of (w/4, h/4).

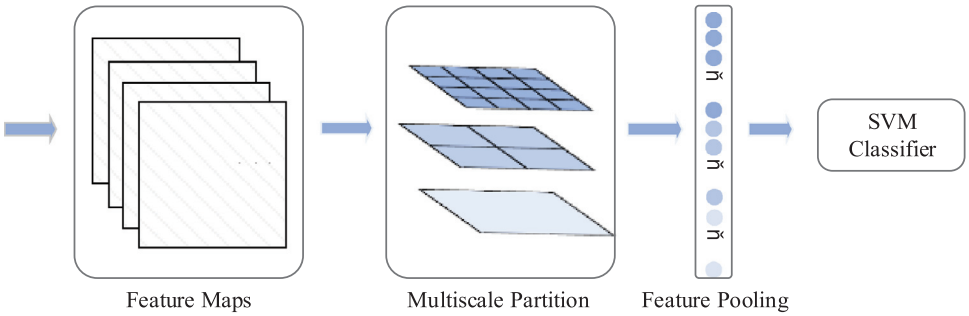

Fig. 3. Flowchart of the spatial pyramid pooling layer (SPP).

(2) In the partition based on the second scale, the same input feature map of the first step is divided into 4 sub-image blocks with a size of ($w/2$, $h/2$).
(3) In the partition based on the third scale, the input feature map does not divide anymore and has the original size of (w, h).

The above SPP layer produces multi-scale spatial features of oceanic eddies, which are considered to have better distinguishing ability than the features learned in single scales. The SPP layer can enhance the robustness of DeepEddy on high dynamic oceanic eddies detection, as will be verified in the experimental section below.

When the invariant features of oceanic eddies are learned hierarchically, a simple linear SVM classifier is chosen to obtain the final results of eddy detection based on these features. It uses the linear kernel $\kappa(x_i, x_j) = x_i^T x_j$ and the hinge loss $Loss = \max(0, 1 - y_i(w^T x_i + b))$.

To sum up, the main parameters of the DeepEddy architecture include the filter size k_1, k_2, the number of convolutional layers, the number of filters in each convolutional layer L_1, L_2, ⋯, L_h, the block size for local histogram, the overlapping ratio between blocks and the pyramid vector.

3. Experimental analysis

In this section, we describe the experiments carried out to verify the detection accuracy of DeepEddy. All experiment were performed on a computer with an Intel i7 CPU with 16GB memory, Windows 10 OS and Matlab 2015b.

3.1. Datasets

We established our dataset starting from the SAR images generated by ENVISAT and ERS-2 Satellites between 2005 and 2011 (provided by European Space Agent https://earth.esa.int). ENVISAT SAR images are C band (5.3 GHz) data with three modes: the image mode, the alternating polarization mode with a swath width of 100 km and approximately 30 m × 30 m spatial resolution, and the wide swath mode with a swath width of over 400 km and 150 m × 150 m resolution. ERS-2 also works in the C band with 30 m × 30 m resolution. Most of the SAR images used in our experiments are in VV polarization mode, which is more suitable for the study of ocean currents and ocean waves. The vertical and horizontal movement of oceanic eddies can cause significant fluctuation of ocean currents and waves in the coverage area. The oceanic eddies training dataset in SAR images has different resolutions and is perfect for the detection of oceanic eddies at different scales.

Although the DeepEddy method is independent of geo-spatial locations and temporal variations, we have to constrain our datasets to a certain location and time to maintain the correct association with oceanic eddies generation. In order to improve the efficiency of manual annotation, we used the SAR images covering the area 13°N to 25°N and 116°E to 125°E, which is where Kuroshio generates several oceanic eddies [32].

The oceanic eddies training datasets used in this study were manually annotated by three trained persons. We annotated a total of 159 oceanic eddies in the specific region from 2005 to 2011 on ENVISAT and ERS-2 SAR images, which vary significantly not only in eddies position, scale, orientation and vorticity but also in visual representation and textures. Most of the annotated oceanic eddies images have various sizes ranging from about 100 × 100 pixels (~3 km) to 650 × 650 pixels (~97.5 km); some are above 1000 × 1000 pixels (over 150 km). We resized the oceanic eddies to the specific size of 280 × 280 pixels, which was determined by choosing the dominant size in the statistical distribution of oceanic eddies image sizes. To train a more generalized and more robust deep learning method, we further enlarged the training dataset by data augmentation. To our knowledge, there are no publicly available remotely sensed image datasets at present. We established a relatively more abundant dataset for future eddies studies, and we will continue to update and enrich this dataset in the future. The dataset is available at https://github.com/YanlingDu-SHOU/Mesoscale-Oceanic-Eddy-Dataset.

3.2. Data augmentation

An adequate and diverse training dataset is the key foundation for achieving an excellent image detection accuracy based on the deep learning method. However, the large volume training dataset would consume much labor, which is hard to achieve in practice. Data augmentation is an essential way to increase the volume and variety of the training dataset in deep learning, and the training dataset is usually enlarged by 100–1000 times [40]. In this study, we use image translational transformation, rotation transformation and noise transformation to enlarge the volume of the training dataset, as shown in Fig. 4, leading to a 125 times augmentation and a total of 20,000 samples.

3.3. Training method

To train DeepEddy, we used the 10-fold cross validation on the dataset. This means that the Eddy Dataset was randomly divided into 10 groups, one of which was used as a test set at one time, and the remaining 9 groups were used as a training set for model training. Then, we trained the DeepEddy method varying the following parameters: the number of filters (the number of filters in the first and second layers were defined as L_1 and L_2, respectively); the size of each filter (k_1 and k_2 are the filter sizes in the first layer and second layer, respectively); and the ratio of block overlap (BOR). In this way, we could investigate the influence that each parameter had on the detection accuracy. In addition, we compared the detection accuracy of PCANet [31] and DeepEddy, using the same parameter settings. In the pooling layer, the histogram block size (B) was fixed to 8 × 8, and the spatial pyramid pooling (SPP) was set to 4 × 4, 2 × 2, and 1 × 1 sub-regions for all of the experiments.

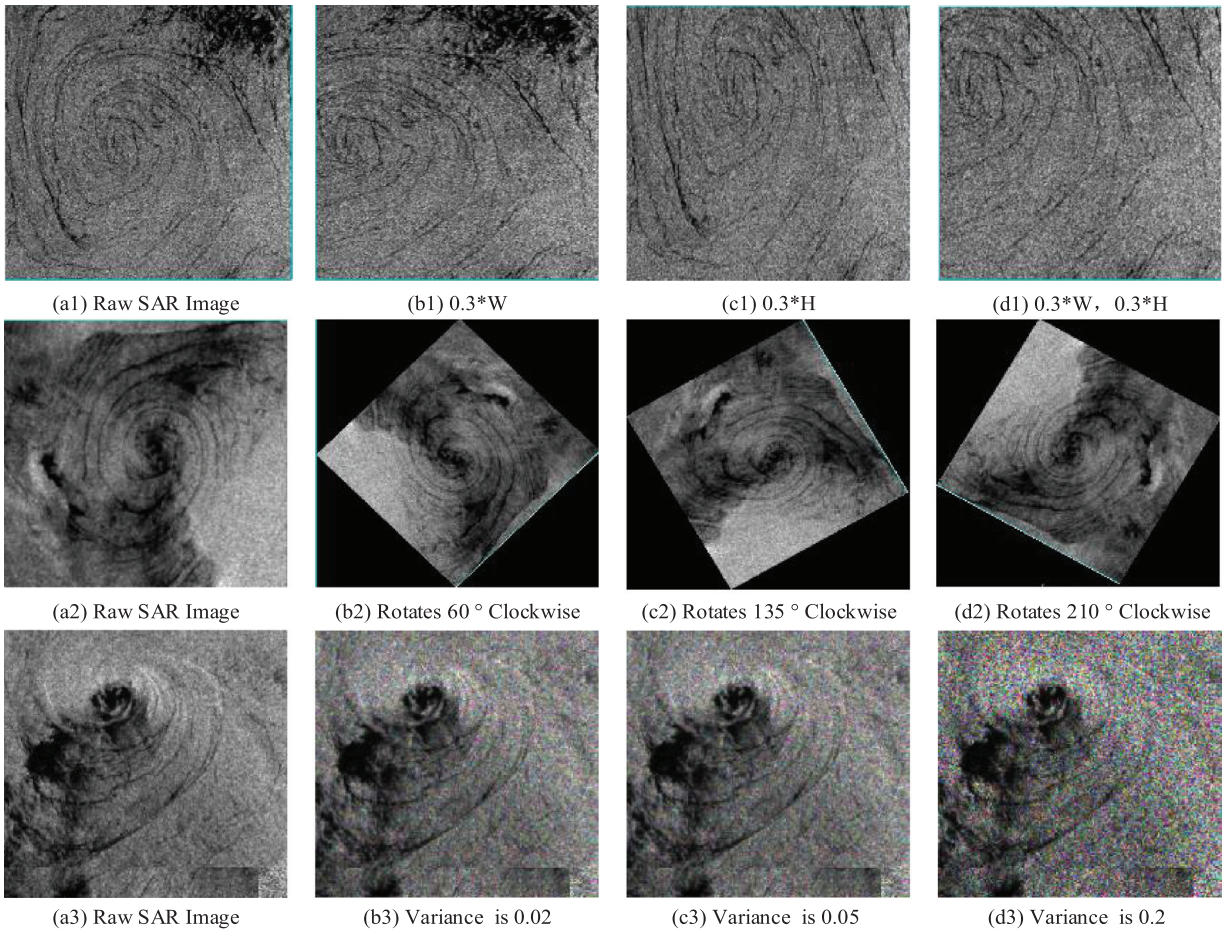

Fig. 4. The oceanic eddies training dataset is enlarged based on data augmentation. (a1)–(d1) are example translational transformations. (a2)–(d2) are example rotation transformations. (a3)–(d3) are example noise transformations.

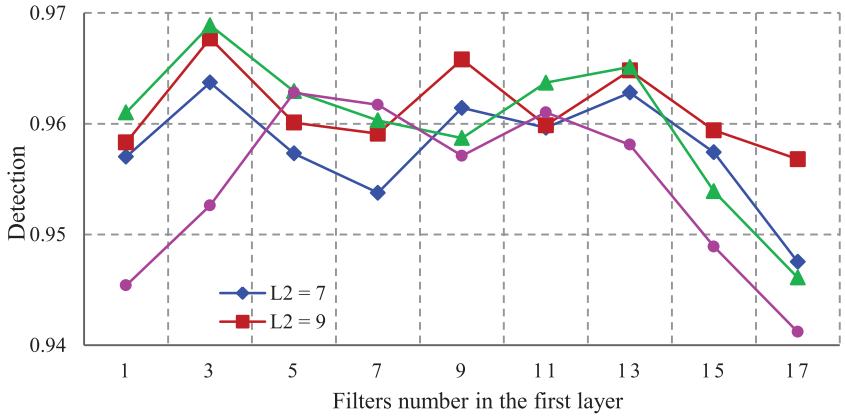

Fig. 5. Oceanic eddy detection accuracy under different numbers of filters. Given $L_2 = 7$, 9, 11 and 13, the changes in the detection accuracy with varying the number of filters in the first layer (L_1) at an interval of 2.

3.4. Impact of the number of filters

To study the impact that the number of filters has on eddy detection accuracy, we varied in turns the number of filters in each layer of DeepEddy, while keeping the other layers constant. Firstly, we varied the number of filters L_1 in the first layer from 1 to 17, at an interval of 2, given the fixed number of filters L_2 in the second layer. The filter sizes of DeepEddy were $k_1 = k_2 = 5$, and the overlap ratio was 0.5. Fig. 5, shows the changes in detection accuracy with L_1 when L_2 was set to 7, 9, 11 and 13, respectively. The model achieves the best detection accuracy at $L_1 = 3$ when L_2 was fixed at 7, 9 and 11, and $L_1 = 5$ when $L_2 = 13$. The results show that the number of filter layers does impact the detection accuracy, and the numbers in the two layers interact with each other. In our experiments, the best accuracy of $96.8 \pm 2\%$ was achieved with a combination of $L_1 = 3$ and $L_2 = 11$. From Fig. 5, there is a declining trend when L_1 is getting bigger (e.g., $L_1 > 13$), and overall, when $L_2 = 7$ and 13, the model achieves lower accuracy than when $L_2 = 9$ and 11. According to the trends, it is suggested that the number of filters in the

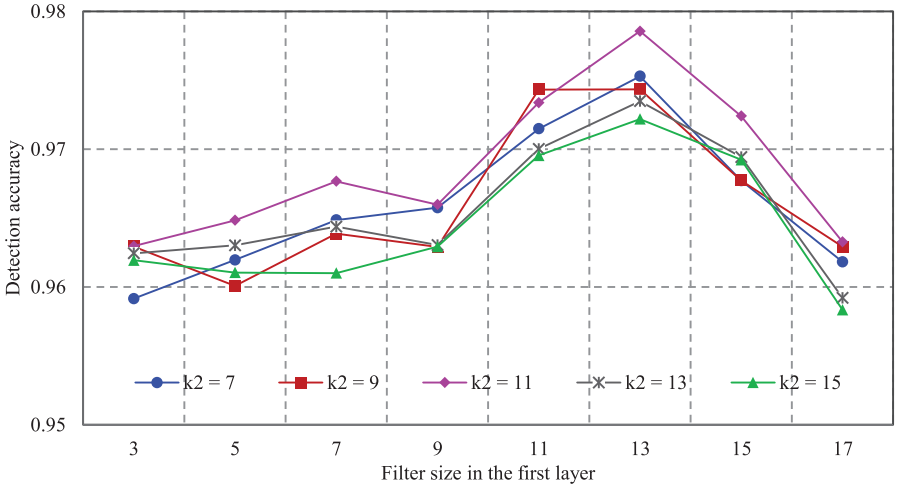

Fig. 6. Oceanic eddy detection accuracy under different filter sizes.

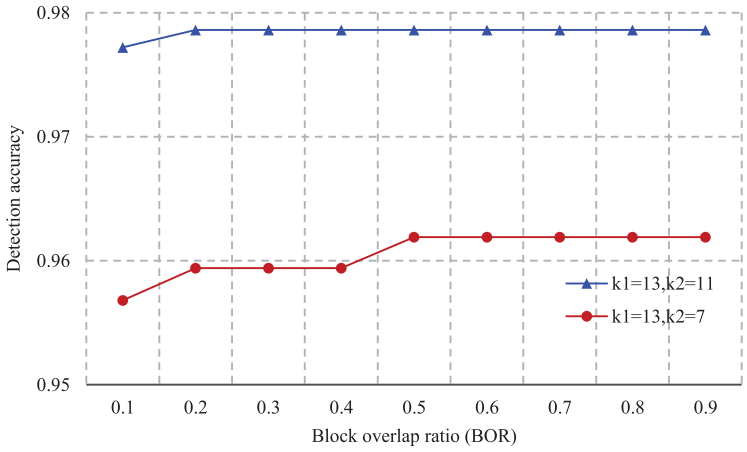

Fig. 7. Oceanic eddy detection accuracy under different block overlap ratios (BORs).

first layer should be relatively small, whereas the number of filters in the second layer should be larger than that of the first layer but should not be too large either.

3.5. Impact of the filter size

The impact of filter size of DeepEddy was further investigated by changing the filter size in both convolutional layers. In particular, we set the second layer filter size $k_2 = 7$, 9, 11, 13 and 15 when changing the first layer filter size k_1 from 1 to 17 with an interval of 2. The filter numbers were set to $L_1 = 3$ and $L_2 = 11$ based on the results in Section 3.5. The block overlap ratio (BOR) was 0.5. As the results show in Fig. 6, the accuracy of DeepEddy increases with the increase in the filter size in the first layer until $k_1 = 13$ and then decreases. k_2 plays little part in this case. The best detection accuracy of DeepEddy was $97.8 \pm 1\%$, with $k_1 = 13$ and $k_1 = 11$.

3.6. Impact of the block overlap ratio

Given the best number of filters ($L_1 = 3$, $L_2 = 11$), the best filter sizes ($k_1 = 13$, $k_2 = 11$), and the second-best filter sizes ($k_1 = 13$, $k_2 = 7$), we changed the block overlap ratio (BOR) between 0.1 and 0.7 (at 0.1 steps), achieving the results depicted in Fig. 7. The detection accuracy tends to be stable with the increase in the block overlap ratio when $k_1 = 13$, $k_2 = 11$. Meanwhile, the detection accuracy has a small step-by-step increase when $k_1 = 13$ and $k_2 = 7$ and becomes stable when BOR $>= 0.5$. This suggests that, as long as the BOR is greater than 0.5, it has no effect on the accuracy. In the case of the above parameter settings, the best accuracy was $97.86 \pm 0.7\%$, which is considerably higher than the best accuracy of 80% gained in traditional methods [9].

3.7. Intermediate outputs of oceanic eddies in SAR images of DeepEddy

To gain a better understanding of the DeepEddy procedure, we visualize here some of its intermediate outputs. These include the convolutional layer filters and the output features of the second convolutional layer, under the DeepEddy parameters that have resulted in a 97.86% detection accuracy, that is: $k_1 = 13$ and $k_2 = 11$, $L_1 = 3$ and $L_2 = 11$, BOR = 0.5. Fig. 8(a) shows the filters learned by the two stages of PCA. These filters with different patterns can extract the most important features of the input image from different angles. Fig. 8(b) shows the 33 feature maps extracted from an example eddy in the second PCA convolutional layer. These feature maps illustrate some spiral shape and texture features in different levels of granularity and grayscale. In DeepEddy, the filters are learned and feature extraction are fully automatic, without requiring the costly intervention of domain experts.

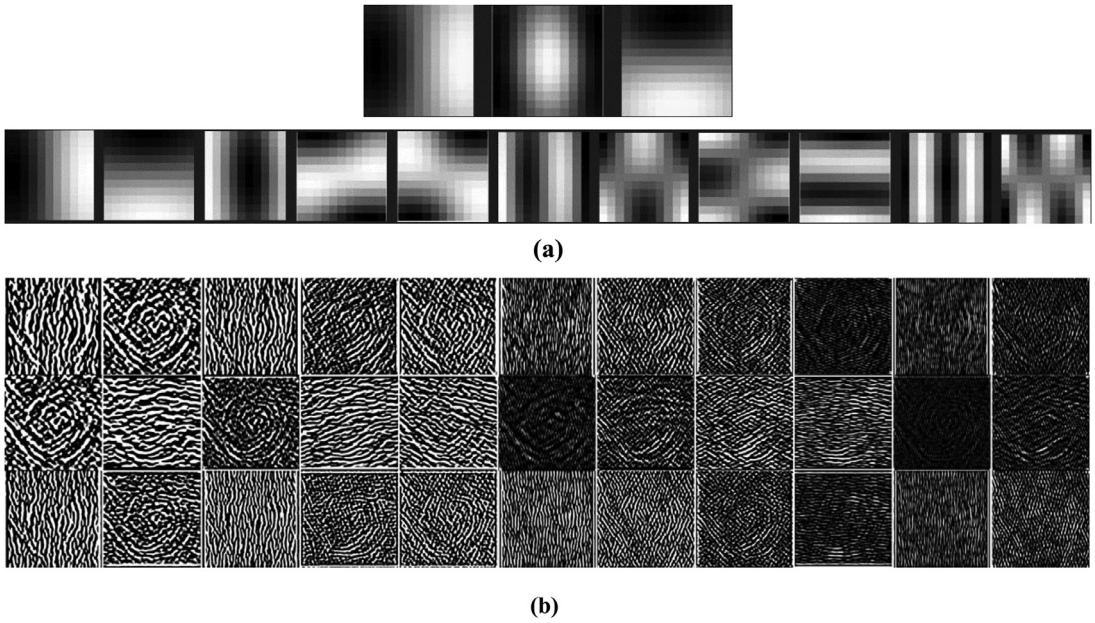

Fig. 8. Visualization of intermediate outputs of DeepEddy. **(a)** The 3 filters learned by the first PCA stage with the size of 13×13 in the top row, and the 11 filters learned by the second PCA stage with the size of 11×11 in the bottom row. **(b)** The 33 output features of the second PCA convolutional layer with the size of 280×280.

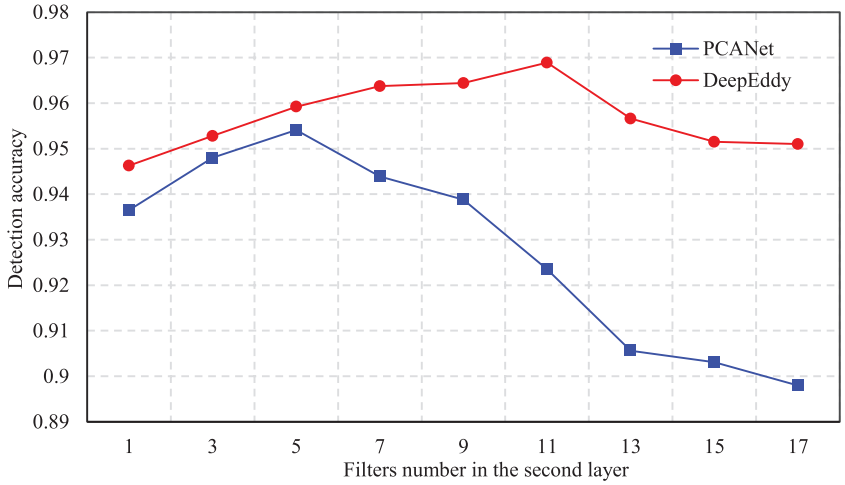

Fig. 9. Comparison of PCANet and DeepEddy detection accuracy with a different number of filters. Given that the filter number in the first layer equals 3, the accuracy changes with an increase in the filter number in the second layer from 1 to 17.

4. Comparative evaluations

4.1. Performance of the detection accuracy of PCANet and DeepEddy

4.1.1. Comparison considering the number of filters

First, we compared the detection accuracy of PCANet and DeepEddy with different filter numbers. The parameters of the filter sizes and overlap ratio of both methods were set to $k_1=13$, $k_2=11$ and 0.5, respectively. The number of filters in the first layer was fixed to $L_1=3$, and the number of filters in the second layer was gradually increased from 1 to 17 with an interval of 2. As illustrated in Fig. 9, the model detection accuracy of DeepEddy is slightly better than that of PCANet when L_2 is less than 5 but significantly better when L_2 is more than 5. In addition, the difference between the two methods gets larger as L_2 increases, which proves that DeepEddy has stronger robustness and higher accuracy.

4.1.2. Comparison considering filter size

Then, we explored the model detection accuracy of the two methods with different filter sizes. In the following four group experiments, the parameters of the filter number and overlap ratio of both methods were set to $L_1=3$, $L_2=11$ and 0.5. The filter sizes in the second layer were set to 3, 5, 9, and 15, and the filter sizes were varied in the first layer gradually from 3 to 17 with an interval of 2. The results shown in Fig. 10(a)–(d), account for a better performance of DeepEddy. In particular, with the gradual increase in the filter size in the first layer, the detection accuracy of DeepEddy improves more rapidly than PCANet. The results verify the validity of adding SPP in the architecture, which can extract the multi-scale spatial features of oceanic eddies and enhances the robustness of DeepEddy.

4.2. Comparisons with non-deep learning methods

At present, automatic detection methods are mainly based on SST and SSH remote sensing data [9–29]. It is inappropriate to compare

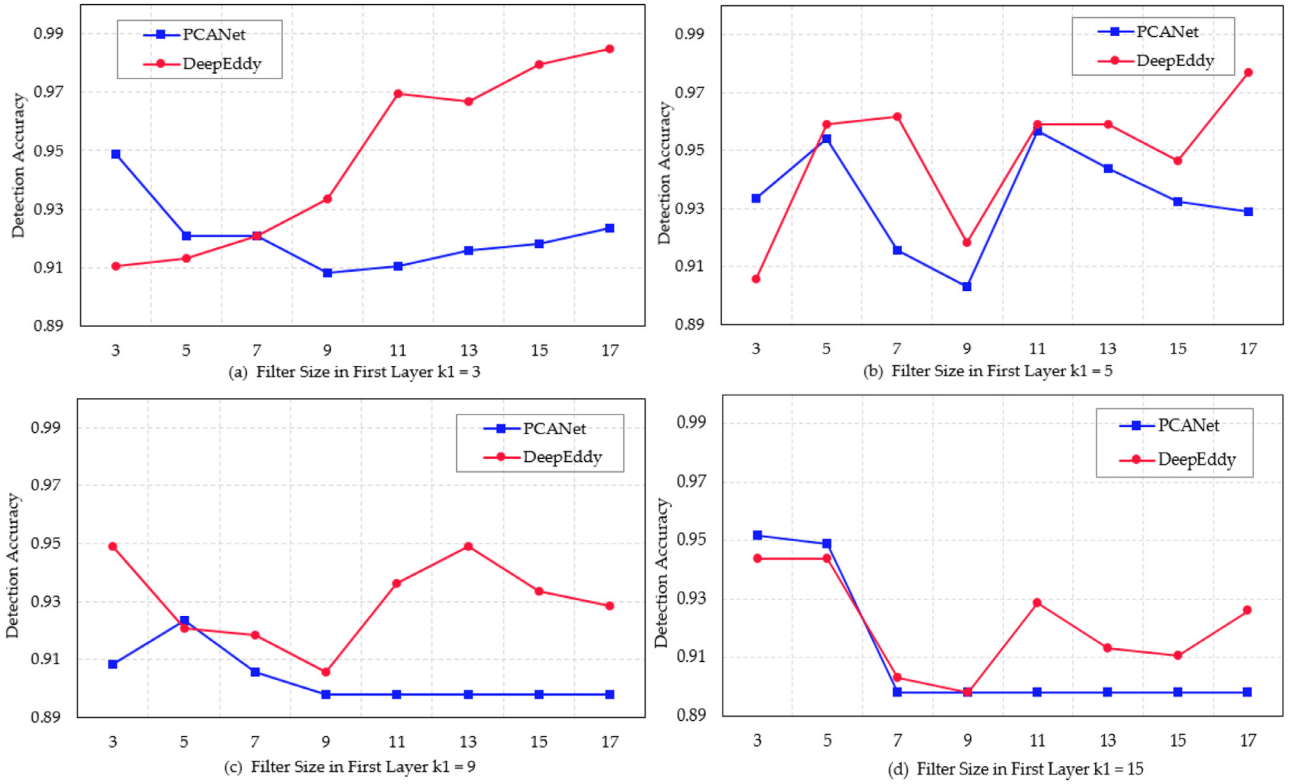
PCANet
DeepEddy
Detection Accuracy
(a) Filter Size in First Layer k1 = 3
(b) Filter Size in First Layer k1 = 5
(c) Filter Size in First Layer k1 = 9
(d) Filter Size in First Layer k1 = 15

Fig. 10. Comparison of PCANet and DeepEddy detection accuracy with different sizes of filters. The filter sizes in the second layer are set to (a) 3, (b) 5, (c) 9, and (d) 15, and the filter size is varied in the first layer gradually from 3 to 17 with an interval of 2.

altimetry-based methods with DeepEddy due to the different data source types and different spatial scales of eddies considered. SAR image-based oceanic eddy detection methods mostly use visual interpretation [28,31,32,41], which relies on some specific visual features [41] to identify the oceanic eddies by eye observation on SAR images. An automatic classifier with SAR image [9] claimed 80% accuracy for several ocean mesoscale atmospheric phenomena, but unfortunately, oceanic eddies were not included in that study. Therefore, we compared our method with a visual interpretation method in SAR by Xu and Yang [32].

Without the exact eddy ground truth in the SAR images from [32], we could not discuss the method accuracy comparison with Xu and Yang's. Here, we only compared our eddy detection results with those presented in [32], which applied visual interpretation based on the visual features of eddies [41]. According to the study in [32], they used the SAR images at Kuroshio region and from the following observational data: ERS-2 and ENVISAT SAR images from 2005 to 2011 with three modes; the image mode (IM), the alternating polarization mode (AP); and the wide swath mode (WS). From those images, 426 frames were first selected manually, and then a total of 60 distinct eddies were identified in 37 SAR images. From the same SAR images described in [32], DeepEddy automatically detected 104 distinct eddies in 65 SAR images, among which there are 42 WS mode, 16 IM mode and 7 AP mode of SAR images. The number of oceanic eddies by DeepEddy is significantly higher than those identified by the visual interpretation method in [32]. The radius distributions of the oceanic eddies detected by both methods are shown in Fig. 11.

Most of the eddies detected by both methods have a radius of 0–21 km, accounting for 88.5% and 93% of all eddies identified by DeepEddy and the visual interpretation method [32], respectively. However, the number of eddies detected by the two methods is significantly different in every radius interval. The number of eddies detected by DeepEddy in each radius interval is generally more than the visual interpretation method. Many eddies with a radius of 0–12 km are expected

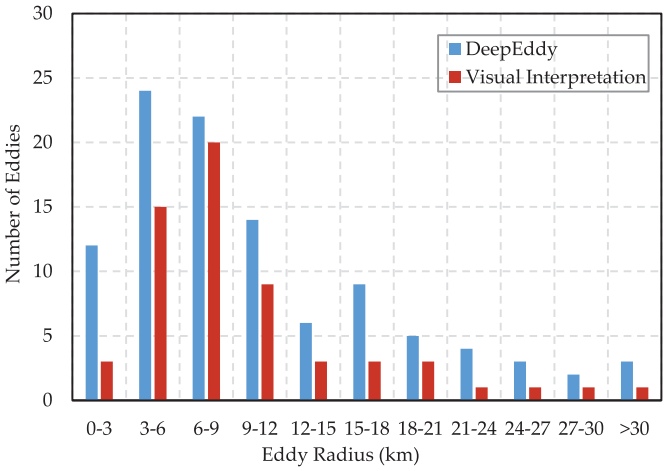

Fig. 11. Comparison of the radius distribution of eddies detected by DeepEddy and the visual interpretation method in the area 13°N to 25°N and 116°E to 125°E where Kuroshio passes through via ENVISAT and ERS-1/2 SAR images from 2005 to 2011.

because small-scale oceanic eddies are mainly related to the complex terrain in the studied Kuroshio sea area.

We further compared the eddy detection results of the two methods based on the inter-annual variations from 2005 to 2011, as shown in Fig. 12. Two observations can be made. Firstly, the number of oceanic eddies detected in SAR images shows large inter-annual variations. The eddies detected in 2007 are many more than those in other years. This phenomenon could be explained by the fact that the generation of small-scale oceanic eddies is closely related to the wind field, temperature field and topographic conditions of the local area. There were El Niño and

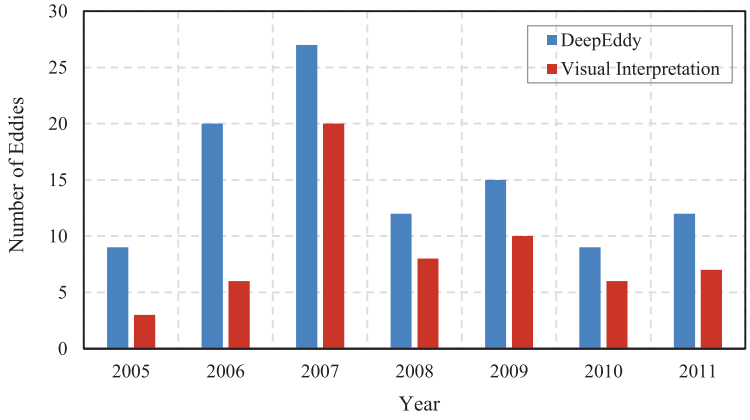

Fig. 12. Yearly distribution of eddies detected in SAR images in the area 13°N to 25°N and 116°E to 125°E between 2005 to 2011. The blue bar and red bar are the results detected in SAR images by DeepEddy and the visual interpretation method, respectively. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

La Niña events that occurred in 2007. Secondly, the inter-annual variation trend of the oceanic eddies detected by DeepEddy is substantially the same as that of the visual interpretation method, except for in 2006. The exception is possible due to a small number of SAR images manually selected in 2006 in [32]. In addition, the visual interpretation method is subjective so that the results might have been affected by one's knowledge and interpretation ability.

In summary, DeepEddy achieves very similar trends to the visual interoperation method in terms of the radius and inter-annual distributions of eddy detection results. It indirectly demonstrates the correctness of DeepEddy. On the other hand, DeepEddy detected more eddies than the visual interoperation method. It shows the effectiveness of DeepEddy compared to a manual process.

5. Conclusions

In this paper, we have studied oceanic eddy detection methods and proposed a DeepEddy architecture based on a simple deep network based on the PCA filter convolution neural network. DeepEddy can learn the high-level and invariant features of oceanic eddies, providing a fully automatic and accurate eddy detection.

Our experiments on the oceanic eddies dataset have demonstrated that the DeepEddy can reach a detection accuracy of $97.8 \pm 1\%$. Regarding the investigation of DeepEddy parameters on detection accuracy, we could draw a number of conclusions of the parameter settings: the filter size should not be too small in order to contain enough local information of the eddies; increasing the block overlap ratio would not have much impact after a certain ratio such as 0.5; the number of filters in different convolutional layers significantly impacted the accuracy. The paramters' impacts on the dection accuracy can serve as a reference for future work. Additionally, the comparison of PCANet and DeepEddy detection accuracy shows that the SSP architecture for multi-scale feature fusion improves detection accuracy due to its robustness in various structures and scales of eddies.

What makes DeepEddy particularly appealing is its use of a light learning strategy, unlike the typical deeper learning networks that have more sophisticated strategies and train millions of parameters. The DeepEddy learning architecture allows to perform well even on a relatively small training datasets, which enables it to be easily extended to other detection tasks using other remotely sensed data. The comparison of oceanic eddy detection results show that DeepEddy can detect more eddies than the visual interpretation method, mainly because DeepEddy can detect the incomplete oceanic eddies as well as large-scale eddies.

In the future, we will further enrich the Eddy dataset, to include different regions of the ocean and annotate the cyclonic and anticyclonic eddies separately in the training dataset building process, which will allow for better differentiation. In addition, we would like to research the ability of DeepEddy on multisource remotely sensed data and enrich the Eddy dataset with other remotely sensed data.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (NSFC) Grant [41671431]; the Program for Professor of Special Appointment (Eastern Scholar) at Shanghai Institutions of Higher Learning [TP2016038]; and the Capacity Development for Shanghai Local College Project Grant by Science and Technology Commission of Shanghai Municipality [15590501900 & 17050501900].

Authors' contributions

Yanling Du designed the oceanic eddy detection method, and together with Wei Song, performed the experiments and result analysis and wrote the paper. Yanling Du, Wei Song and Qi He established the oceanic eddy training dataset. Qi He, Dongmei Huang, Antonio Liotta and Cheng Su provided scientific advice on the method and analysis of results and revised the overall manuscript.

Conflicts of interest

The authors declare no conflict of interest.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.inffus.2018.09.006.

References

[1] M.Z. Uddin, M.M. Hassan, A. Almogren, A. Alamri, M. Alrubaian, G. Fortino, Facial expression recognition utilizing local direction-based robust features and deep belief network, IEEE Access 5 (2017) 4525–4536.
[2] A. Augimeri, G. Fortino, M.R. Rege, V. Handziski, A. Wolisz, A cooperative approach for handshake detection based on body sensor networks, in: 2010 IEEE International Conference on Systems Man and Cybernetics (SMC), 2001, pp. 281–288.
[3] Y.T. Chan, S.J. Wang, C.H. Tsai, Real-time foreground detection approach based on adaptive ensemble learning with arbitrary algorithms for changing environments, Inf. Fusion 39 (2018) 154–167.
[4] P.Y. Le Traon, R. Morrow, Ocean currents and eddies, Int. Geophys. 69 (2001) 171–215.
[5] L.L. Fu, D.B. Chelton, P.Y. Le Traon, R. Morrow, Eddy dynamics from satellite altimetry, Oceanography 23 (4) (2010) 14–25.
[6] J.H. Faghmous, M. Le, M. Uluyol, V. Kumar, S. Chatterjee, A parameter-free spatio-temporal pattern mining model to catalog global ocean dynamics, in: 13th IEEE International Conference on Data Mining (ICDM), 2013, pp. 151–160.

[7] W. Alpers, P. Brandt, A. Lazar, D. Dagorne, B. Sow, S. Faye, P. Brehmer, A small-scale oceanic eddy off the coast of West Africa studied by multi-sensor satellite and surface drifter data, Remote Sens. Environ. 129 (2013) 132–143.
[8] J. Marcello, F. Eugenio, S. Estrada-Allis, P. Sangrà, Segmentation and tracking of anticyclonic eddies during a submarine volcanic eruption using ocean colour imagery, Sensors 15 (4) (2015) 8732–8748.
[9] K. Topouzelis, D. Kitsiou, Detection and classification of mesoscale atmospheric phenomena above sea in SAR imagery, Remote Sens. Environ. 160 (2015) 263–272.
[10] L.L. Fu, R. Ferrari, Observing oceanic submesoscale processes from space, Eos Trans. Am. Geophys. Union 89 (48) (2008) 488-488.
[11] J. Yi, Y. Du, C. Zhou, F. Liang, M. Yuan, Automatic identification of oceanic multieddy structures from satellite altimeter datasets, IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 8 (4) (2015) 1555–1563.
[12] C. Dong, F. Nencioli, Y. Liu, J.C. McWilliams, An automated approach to detect oceanic eddies from satellite remotely sensed sea surface temperature data, IEEE Geosci. Remote Sens. Lett. 8 (6) (2011) 1055–1059.
[13] C. Zhang, H. Li, S. Liu, L. Shao, Z. Zhao, H. Liu, Automatic detection of oceanic eddies in reanalyzed SST images and its application in the East China Sea, Sci. China Earth Sci. 58 (12) (2015) 2249–2259.
[14] S.A. Henson, A.C. Thomas, A census of oceanic anticyclonic eddies in the Gulf of Alaska, Deep Sea Res. I 55 (2) (2008) 163–176.
[15] J. Isern-Fontanet, E. García-Ladona, J. Font, Identification of marine eddies from altimetric maps, J. Atmos. Oceanic Technol. 20 (5) (2003) 772–778.
[16] J. Isern-Fontanet, E. García-Ladona, J. Font, Vortices of the Mediterranean Sea: an altimetric perspective, J. Phys. Oceanogr. 36 (1) (2006) 87–103.
[17] D.B. Chelton, M.G. Schlax, R.M. Samelson, R.A. de Szoeke, Global observations of large oceanic eddies, Geophys. Res. Lett. 34 (15) (2007).
[18] D. Matsuoka, F. Araki, Y. Inoue, H. Sasaki, A new approach to ocean eddy detection, tracking, and event visualization – application to the northwest Pacific Ocean, International Conference on Computational Science, 2016 11–13 June, Wuhan, China.
[19] D.B. Chelton, M.G. Schlax, R.M. Samelson, Global observations of nonlinear mesoscale eddies, Prog. Oceanogr. 91 (2) (2011) 167–216.
[20] F. Nencioli, C. Dong, T. Dickey, L. Washburn, J.C. McWilliams, A vector geometry–based eddy detection algorithm and its application to a high-resolution numerical model product and high-frequency radar surface velocities in the Southern California bight, J. Atmos. Oceanic Technol. 27 (3) (2010) 564–579.
[21] S. Williams, M. Hecht, M. Petersen, R. Strelitz, M. Maltrud, J. Ahrens, B. Hamann, Visualization and analysis of eddies in a global ocean simulation, Comput. Graphics Forum 30 (3) (2011) 991–1000.
[22] A. Chaigneau, A. Gizolme, C. Grados, Mesoscale eddies off Peru in altimeter records: identification algorithms and eddy spatio-temporal patterns, Prog. Oceanogr. 79 (2) (2008) 106–119.
[23] Y. Lai, H. Zhou, J. Yang, Y. Zeng, B. Wen, Submesoscale eddies in the taiwan strait observed by high-frequency radars: detection algorithms and eddy properties, J. Atmos. Oceanic Technol. 34 (4) (2017).
[24] A. Turiel, J. Isern-Fontanet, E. García-Ladona, Wavelet filtering to extract coherent vortices from altimetric data, J. Atmos. Oceanic Technol. 24 (12) (2007) 2103–2119.
[25] A.M. Doglioli, B. Blanke, S. Speich, G. Lapeyre, Tracking coherent structures in a regional ocean model with wavelet analysis: application to Cape Basin eddies, J. Geophys. Res. 112 (C5) (2007).
[26] Y. Chamber, A. Garg, V. Mithal, I. Brugere, V. Kumar, M. Lau, S. Boriah, A novel time series based approach to detect gradual vegetation changes in forests, in: Conference on Intelligent Data Understanding (CIDU), 2011, pp. 248–262.
[27] J.H. Faghmous, Y. Chamber, S. Boriah, F. Vikebø, S. Liess, M. dos Santos Mesquita, V. Kumar, A novel and scalable spatio-temporal technique for ocean eddy monitoring, Conference on Artificial Intelligence (AAAI), 2012.
[28] S. Karimova, M. Gade, Sub-mesoscale eddies seen by spaceborne radar, in: Proc. EMEC 10-MEDCOAST 2013, vol. 1, 2013, pp. 665–676.
[29] A.Y. Ivanov, A.I. Ginzburg, Oceanic eddies in synthetic aperture radar images, J. Earth Syst. Sci. 111 (3) (2002) 281.
[30] P.M. DiGiacomo, B. Holt, Satellite observations of small coastal ocean eddies in the Southern California bight, J. Geophys. Res. 106 (C10) (2001) 22521–22543.
[31] S. Karimova, Spiral eddies in the Baltic, Black and Caspian Seas as seen by satellite radar data, Adv. Space Res. 50 (8) (2012) 1107–1124.
[32] G. Xu, J. Yang, C. Dong, D. Chen, J. Wang, Statistical study of submesoscale eddies identified from synthetic aperture radar images in the Luzon strait and adjacent seas, Int. J. Remote Sens. 36 (18) (2015) 4621–4631.
[33] Franz, K., Roscher, R., Milioto, A., Wenzel, S., & Kusche, J. (2018). Ocean eddy identification and tracking using neural networks. arXiv:1803.07436v2 [cs.CV], May 2018.
[34] A. Krizhevsky, I. Sutskever, G.E. Hinton, Imagenet classification with deep convolutional neural networks, in: Advances in Neural Information Processing Systems, 2012, pp. 1097–1105.
[35] T.H. Chan, K. Jia, S. Gao, J. Lu, Z. Zeng, Y. Ma, PCANet: a simple deep learning baseline for image classification? IEEE Trans. Image Process. 24 (12) (2015) 5017–5032.
[36] D.C. Mocanu, E. Mocanu, P.H. Nguyen, M. Gibescu, A. Liotta, A topological insight into restricted Boltzmann machines, Mach. Learn. 104 (2-3) (2016) 243–270.
[37] D. Huang, Y. Du, Q. He, W. Song, K. Liu, A. Liotta, Deepeddy: a simple deep architecture for mesoscale oceanic eddy detection in SAR images, 14th IEEE International Conference on Networking, Sensing and Control (ICNSC), 2017 Calabria, Southern Italy, May 16–18.
[38] H.R. Shahdoosti, H. Ghassemian, Combining the spectral PCA and spatial PCA fusion methods by an optimal filter, Inf. Fusion 27 (C) (2016) 150–160.
[39] H. Ghassemian, A review of remote sensing image fusion methods, Inf. Fusion 32 (PA) (2016) 75–89.
[40] Z. Gan, R. Henao, D. Carlson, L. Carin, Learning deep sigmoid belief networks with data augmentation, Artif. Intell. Stat. (2015) 268–276.
[41] P.M. DiGiacomo, B. Holt, Satellite observations of small coastal ocean eddies in the Southern California bight, J. Geophys. Res. 106 (C10) (2001) 22521–22543.